From volatility smiles to the volatility of volatility

Published version:

DOI:10.1007/s10203-019-00263-w

From volatility smiles to the volatility of volatility*

Bernard Dumas
INSEAD and University of Torino AXA Chair

Elisa Luciano
University of Torino and Collegio Carlo Alberto, Italy

September 10, 2018, revised June 18, 2019

Abstract

The paper reviews models of the option surface and reduced-form models for stochastic volatility in continuous time, under the risk-neutral measure. It defines "forward volatilities," analogous to forward interest rates in the theory of the term structure, and provides a proof that the forward volatility is a conditional expected value, under the risk-neutral measure, of the future spot volatility. The theory developed here is the analog of Heath-Jarrow-Morton bond-pricing theory. The link is established between forward volatilities and so-called "model-free" volatility measures such as the VIX.

KEYWORDS: stochastic volatility, implicit volatility, forward volatility, VIX

JEL classification numbers: G13, G17

*Part of this paper is published as Chapter 17 in Bernard Dumas and Elisa Luciano, The Economics of Continuous-Time Finance, pp. 437-453.

The Black-Scholes option-pricing formula is based on the assumption that the volatility of the underlying security is constant. By inversion of the formula, for a given value of the underlying, the value of any option provides an "implied volatility" σ_I. If the assumption were correct, the very same value σ_I would be obtained from the prices of options of all maturities and strikes. Empirically, it is not.

The standard empirical phenomenon that questions the validity of the Black-Scholes model as a description of the actual dynamics of asset prices is the so-called option "smile." Far from being constant, empirical observations for a given maturity exhibit higher implied volatility for far in-the-money and far-out-of-the-money options, smaller volatility for close-to-the-money options. Implied volatility as a function of moneyness is therefore like a smile, when symmetric over the strikes, or more often a "smirk," when asymmetric. And the volatility surface is made by as many smiles or smirks as maturities.

A somewhat related phenomenon is the so-called "leverage effect," which consists in a negative correlation between market returns and changes in volatility. Empirically, volatility usually rises while market prices are low or volatility is low when the market is high. One intuitive explanation for that phenomenon is the following. When equity prices rise, firms' leverage, everything else equal, drops, their stocks become less risky and the volatility of equity returns is lower. In empirical tests the numbers do not match this leverage explanation, but the label "leverage effect" serves to describe the negative correlation between volatility and stock-price level.

How can we explain the empirical phenomena that are inconsistent with Black-Scholes modelling? In principle, we could construct a structural model for volatility. In order to do this, we should have good theoretical foundations for the way in which volatility evolves over time in a financial market. This obviously depends on the way information arrives and on what type of information arrives, public *vs.* private. Indeed, roughly speaking, "purely public" information arriving repeatedly generates unconditional volatility but no trading volume.[1] "Purely private" information generates volume but no change in price.

The resulting behavior of volatility would likely be stochastic and, above all, highly model specific. As a shortcut, one can consider reduced-form models that do not model information processing explicitly, but postulate a dynamics for volatility. The literature tried to do that first while preserving market completeness, then building more flexible models, where market completeness is abandoned. We mention here the continuous-time ones.

Market completeness, which is guaranteed in Black-Scholes, is very practical, in that it gives unique prices and easy one-to-one mappings of relevant variables. So convenient that most models assume it. In reduced-form volatility modelling in continuous time, completeness is safeguarded if spot, or instanta-

[1] If more information arrives into a market and gets incorporated into prices, the conditional volatility is reduced when the information arrives (uncertainty is reduced at that point since consumers learn). But, as this happens day in and day out, there is more movement in price: unconditional volatility increases.

neous, volatility evolves as a deterministic function of the underlying price and time (that is a Deterministic Volatility Function, or DVF), as in the Constant-Elasticity of Variance (CEV) model for stocks or the Cox, Ingersoll and Ross (1985a, b) model for interest rates. Rubinstein (1994) provided a general DVF formulation on a binomial tree while Dupire (1994) and Derman and Kani (1994) provided one in continuous time and defined for the DVF case a concept called "local volatility," its value being calculated from the surface of option prices of different maturities and strikes, just like forward interest rates are calculated from the string of bond prices of different maturities. However, as shown for instance in Dumas *et al.* (1998), DVFs are not sufficiently rich to capture the empirical behavior of volatility and, consequently, to hedge it. Hobson and Rogers (1998) provided a more flexible deterministic volatility function, which is little investigated empirically.

Completeness is maintained also if volatility is stochastic, but depends on the same Wiener risks that affect prices. While this modelling avenue is easy, it has been little studied.

Incompleteness arises in volatility modelling when volatility is stochastic and depends also on a separate Wiener process. This was the approach taken by spot volatility models such as the famous Hull and White (1987) and Heston (1993) models. These models, despite their success, pose some problems of calibration, when trying to capture the "market price of volatility risk," which is the excess return, over and beyond the risk-free rate, due to the presence of volatility as a state variable. The concerns of Dumas *et al.* (1998) are valid also for stochastic-volatility models.

How did the literature cope with incompleteness? The answer is: by restoring completeness, in a fashion analogous to what Heath, Jarrow and Morton (1992) did for the term structure of interest rate. In a nutshell, it filled the degrees of freedom of stochastic volatility modelling with the most strict adherence to the hints on volatility provided by the prices of a continuum of derivatives. There was a key observation in that sense, which dates back to Breeden and Litzenberger: the density of the underlying asset obtains from the second derivative of the call price, taken with respect to the strike price. And the density is linked to the instantaneous volatility of a diffusion. For a diffusion backed by more than one Wiener, roughly said, we can assume that the number of traded options is large enough to make the market complete.

However, the path from Breeden-Litzenberger to the elaboration of a stochastic volatility model that preserves completeness is challenging. Contributions along that path include Dumas (1995), Derman and Kani (1997), Carr and Madan (1998) and Britten-Jones and Neuberger (2000). By analogy with the Heath et al. (1982) theory of forward interest rates, Dumas (1995), in an unpublished note, extended to stochastic volatility the local-volatility concept of Dupire, and called it "forward volatility." Derman and Kani (1997), in a Goldman-Sachs technical note, showed that local (or forward) volatility is the expected value (not conditional on the future price of the underlying) of future volatility under a strike-adjusted forward measure. A proof for underlyings that are semimartingales is in Carmona and Nadtochiy (2009).

The main purpose of this short article is to show, as in Dumas (1995, unpublished) that, with a zero rate of interest and under a Markovian assumption, forward volatility is the expected value of future (spot) volatility under the risk neutral measure, conditional on the underlying price hitting a particular value at the maturity date (Theorem 3 below). The term "forward" volatility, as opposed to "local" volatility, serves to remind one of that property, which is not yet well-known. Once the forward volatility is calibrated, the forecast of future volatility can be obtained.

The importance of the forward volatility lies not only in calibration, but also in the construction of the most famous volatility index, the VIX, produced by the CBOE (http://www.cboe.com/micro/vix/vixwhite.pdf). The VIX is related also to the price of a variance swap. Comparing the latter to the actual realized variance, and making a correction for the covariance between realized variance and consumption growth, or market return as a proxy, one can then determine empirically whether volatility risk is priced in the market (over and beyond their consumption-risk content). In their paper suggestively entitled "The Price of a Smile," Buraschi and Jackwerth (2001) took up that task, based on data on stock-market index options. They use the Generalized Method of Moments (GMM) to test that the two are equal. Overwhelmingly, they reject the null hypothesis of zero price. Options on the index are not redundant assets: they allow one to take bets on volatility.[2] This is the sense in which the whole literature goes from riding the smiles to pricing them.

Following the path from the volatility smile to pricing, the paper proceeds as follows: first we give an analytical account of models that preserve market completeness (Section 1.1) and models that do not (Section 1.2). We get to the core studying forward, stochastic-volatility models (Sections 2 and 3), and their empirical consequences (Section 3.1). We explain the relationship with the VIX index in Section 3.2. Section 4 concludes.

1 Pricing a derivative on the basis of the underlying and volatility

1.1 Time-varying volatility with market completeness

The most straightforward way to extend the strict Black-Scholes model and its constant volatility is to make volatility either a deterministic function of time, or a non-deterministic one that would be driven by the same Wiener processes that appear in the asset price processes. For the sake of simplicity, we explain the point with a single (risky) asset, which follows an Itô diffusive process, in a single-Wiener space and its generated filtration.

[2] Stochastic variance can receive a price in the financial market (over and above consumption risk) only if the utility of the representative investor (imagining there exists one) is not a time-additive von Neumann-Morgenstern utility. Volatility is a "delayed risk:" a change in volatility will only have an effect on the process after the immediate period of investment.

Assume that the asset price includes a Deterministic Volatility Function (DVF):

$$\begin{aligned}\frac{dS(t)}{S(t)} &= \mu(t)\,dt + \sigma(S,t)\,dw(t)\,, \\ S(0) &= S_0.\end{aligned}$$

where $\sigma(S,t)$ is a deterministic function of time t and the asset price S, capturing the leverage effect. The volatility at any future point in time is uniquely determined by calendar time and the value attained at that point by the underlying. This formulation allows for variation over time, but preserves market completeness. It suffices to postulate a function $\sigma(S,t)$ with some parameters to be determined, solve the Black-Scholes PDE for the corresponding option price and fit it to the volatility smile. Riding the smile, indeed. The fit serves to determine the parameters of the volatility function, not the implied volatility as a number, as it would do in the Black- Scholes model.

A well-known example for stocks is the so-called Constant-Elasticity of Variance (CEV) model, in which the volatility is proportional to S^{α} with $\alpha < 1$, namely

$$\frac{dS(t)}{S(t)} = \mu(t)\,dt + \sigma S^{\alpha} dw(t)\,,$$

$\sigma \in \mathbb{R}^{++}$. A notable example in the interest rate domain is the model by Cox, Ingersoll and Ross (1985), in which the spot interest rate evolves as

$$\begin{aligned} dr(t) &= \varrho(\bar{r} - r(t))\,dt + \psi\sqrt{r(t)}dw^*(t)\,, && (1)\\ r(0) &= r_0 > 0, && (2)\end{aligned}$$

where ϱ, $\bar{r}$ and $\psi \in \mathbb{R}^{++}$. Since the market is complete and technical conditions are satisfied, the pricing measure is unique and replication of the option, which is a way to demonstrate the Black-Scholes formula, is still possible. While straightforward, the DVF specifications, including the CEV, have unfortunately done a poor job in capturing the behavior of volatility. For instance, Dumas *et al.* (1998), using weekly S&P data over the period June 1988 to December 1993, show that the empirical performance of any deterministic-volatility model is poor. The calibrated parameters are very unstable. Parameters estimated from one cross-section of option prices observed on one day and fitting the smile that day, regularly become invalid just one week later. As a consequence of their lack of robustness, hedging performance of DVF models is poor. Quite paradoxically, the strict Black-Scholes model performs better in-sample and its hedge ratios are more effective than the DVF ones.

The DVF specification has been generalized by Hobson and Rogers (1998). Define the log returns on discounted asset prices $P(t)$, $Z(t) = \ln\left(P(t)\,e^{-rt}\right)$. Introduce the so-called "offset function of order m," $S^{(m)}(t)$, which "sums" the difference between the current return and previous returns $(Z(t-u)\,, u \leq t)$, raised to the power m, and gives exponentially decreasing importance to returns far in the past (by multiplying them with $e^{-\lambda u}$):

$$S^{(m)}(t) = \int_0^\infty \lambda e^{-\lambda u} \left(Z(t) - Z(t-u)\right)^m du.$$

Then use these as arguments of the volatility functions:

$$\sigma\left(t, Z(t), S^{(1)}(t), ...S^{(n)}(t)\right).$$

The volatility of the log-prices Z_t depends on the offsets of the log- price from its exponentially weighted moving averages. In this way, the model allows volatility to depend not just on the level of the stock price, but also on stock *price changes* and their powers, appropriately weighted. For instance, using only the offset function of order one, one has:

$$dZ(t) = dS^{(1)}(t) + \lambda S^{(1)}(t)\, dt.$$

The advantage of this model is that it captures past price changes while preserving a complete-market and allowing an easy derivation of the PDE for option prices. The empirical investigations of this model are not many: however, they show a clear superiority of the model with respect to the Heston model (Section 1.2.2), at least as far as options written on the S&P are concerned. See Platania and Rogers (2005).

Another way of modeling time-varying volatility without abandoning the safe harbor of complete-market models is to assume that the same Wiener process drives the underlying and the volatility level, as in the following equations:

$$\frac{dS}{S} = \mu(t)\, dt + \sigma(S, Y, t)\, dw(t), \tag{3}$$

$$dY = \mu_Y(t)\, dt + \sigma_Y(S, Y, t)\, dw(t), \tag{4}$$

where w is a one-dimensional Wiener process and Y is a (latent) state variable that, together with the underlying asset price, drives the function σ. The other functions have the usual properties that guarantee existence of a solution for the system (3, 4). Volatility is perfectly correlated with the price (negatively or positively, according to the sign of the parameters and functions), so as to be able to accommodate both the smile and leverage effects. With this approach, the market is still complete. The possible empirical failures in capturing the relevant phenomena are not known, since, to our knowledge, this avenue has not been explored empirically. However, the model is a special case of models presented below, which are tested.

1.2 Stochastic volatility and market incompleteness

Much attention has been devoted to extensions of the previous approaches that are known as "stochastic volatility models." In these models, volatility, or the latent variable behind it named Y, entails the introduction of a second Wiener process:

$$\begin{aligned} \frac{dS}{S} &= \mu\left(t\right)dt + \sigma\left(S,Y,t\right)dw\left(t\right), \\ dY &= \mu_Y\left(t\right)dt + \sigma_Y\left(S,Y,t\right)dw\left(t\right), \end{aligned}$$

where w is now a *two-dimensional* Wiener process. These models are flexible and can capture any source of randomness in the volatility that is not caused by movements in the underlying. But they make the market incomplete, and this opens valuation problems. There arise difficulties in capturing the "market price of volatility risk." The latter difficulty will be apparent when we examine some renowned models of stochastic volatility, namely the Hull-and-White and the Heston models. The closed or semi-closed formulas they provide allow calibration. Even then calibration is not trivial.

1.2.1 Hull and White

This section briefly summarizes the Hull-and-White (1987) model and the way we can calibrate it, using options data. We start by supposing that, *under a risk-neutral probability* P^*, the joint stochastic process for a stock price and its volatility is:[3]

$$\begin{aligned} \frac{dS\left(t\right)}{S\left(t\right)} &= rdt + \sigma\left(t\right)dw_1^*\left(t\right), \\ V\left(t\right) &= \sigma^2\left(t\right), \\ dV\left(t\right) &= \mu_Y\left(\sigma\left(t\right),t\right)V\,dt + \sigma_Y\left(\sigma\left(t\right),t\right)V\,dw_2^*\left(t\right), \end{aligned} \tag{5}$$

where w_1^* and w_2^* are two *independent* scalar Wiener processes under P^*, the drift and diffusion coefficients of V depend on a state variable Y which does not include the underlying price S, and the process V cannot take negative values, exactly as in Hull-and-White (1987). In addition, σ is a stochastic process adapted with respect to the filtration generated by (w_1^*, w_2^*) and the riskless rate r is a constant. Observe that the variance process V evolves on its own, independently of w_1^* or S.

Applying Itô's lemma over the interval $[0,T]$, we get:

$$\ln S\left(T\right) - \ln S\left(0\right) = \int_0^T \left[r - \frac{1}{2}\sigma^2\left(t\right)\right]dt + \int_0^T \sigma\left(t\right)dw_1^*\left(t\right). \tag{6}$$

To calibrate the model, we first price options. To do that, we try to establish a connection with Black and Scholes. Let T be the life of a derivative. Introduce a quantity that is usually labeled "realized variance of $\ln S(T)$," i.e. the true realization of the variance of log returns over an interval $[0,T]$, equal to: $\int_0^T \sigma^2\left(t\right)dt$. Define its average $\overline{V}$ as:

$$\overline{V} = \frac{1}{T}\int_0^T \sigma^2\left(t\right)dt.$$

[3] Note that we wrote the coefficients in the V equation in the percentage form.

Suppose for a second that we give ourselves $\overline{V}$ in (6):

$$\ln S(T) - \ln S(0) = \left[r - \frac{1}{2}\overline{V}\right] T + \int_0^T \sigma(t)\, dw_1^*(t).$$

Since σ is adapted and evolves independently of w_1^*, or S for that matter, $\int_0^T \sigma(t)\, dw_1^*(t)$, for a given volatility path, or conditionally on $\overline{V}$, is a weighted sum of normal increments $dw_1^*(t)$ where the weights are unrelated to the increments. For that reason, conditional on $\overline{V}$, log returns on the underlying are normally distributed, with zero expected value and a variance equal to $\overline{V} \times T$. Hence the price of a call option can be written as the expectation conditioned down of a risk-neutral process:

$$C(x, y) = \int_0^\infty \mathbb{E}^* \left\{ [S(T) - \mathcal{K}]^+ \mid \overline{V}, S(0) = x \right\} \times h\left(\overline{V} \middle| V(0) = y\right) d\overline{V},$$

where h is the density of $\overline{V}$. Given the normality of $\ln\left(S(T)\,|\overline{V}\right)$, we also have an explicit formula for $\mathbb{E}\left\{[S(T) - \mathcal{K}]^+ \mid \overline{V}, x\right\}$. It is just the Black-Scholes formula with volatility $\overline{V}$.

Our solution for the Hull-White option is not complete until we obtain the density of the realized volatility $h\left(\overline{V}\middle| V(0) = y\right)$. In their article, Hull and White provide closed-form expressions for the first three moments of $\bar{V}$ when $\mu_Y = 0$, and a differential equation for h.

Still, the model is not completely specified, neither under the risk-neutral nor under the effective measure, unless we calibrate the parameters $\sigma_Y(\sigma, t)$ and $\mu_Y(\sigma, t)$, which drive the volatility process under the risk-neutral measure, and unless we provide an estimate of the volatility-risk premium, which represents the adjustment from μ_Y to an effective measure. The two parameters $\sigma_Y(\sigma, t)$ and $\mu_Y(\sigma, t)$ should be backed out from option prices. This is usually done by minimizing the so-called pricing error, i.e. the distance between theoretical and actual prices. Distance can be measured by taking the mean square error of the pricing errors, their mean absolute value, or other measures. Given that the Hull-White model, because of the density h, does not provide fully closed-form formulas, computing the pricing error is computationally demanding. The volatility risk premium is even more challenging to obtain. If we want to utilize the Hull-and-White model as is, and switch from the risk-neutral to the effective measure without changing the calibrated drift, we need to assume that this covariance is zero or, in other words, that volatility risk is not priced. This is the main restriction of the model: there is volatility "risk," in the sense that volatility is stochastic, with an additional Wiener process entering the picture, but any estimation of the parameters of the model valid under both measures requires that investors give a price of zero the new source of risk.

1.2.2 Heston

The Heston (1993) model is another specification of stochastic volatility, which reflects mean reversion in volatility, is able to reproduce the leverage effect and permits calibration. Again, we introduce the model, specify its calibration possibilities, and conjecture the form of volatility risk premia.

Under a risk-neutral measure P^*, as above, let the stock and variance processes be

$$\begin{aligned} \frac{dS(t)}{S(t)} &= rdt + \sqrt{V(t)}dw_1^*(t), \\ dV(t) &= \{\varrho \times [\theta - V(t)] - kV(t)\}dt + k\sqrt{V(t)}dw_2^*(t), \end{aligned} \tag{7}$$

where r is a constant rate of interest, $V(0) > 0$ and the parameters satisfy the Feller condition $2\rho\theta > k^2$,so that V stays positive at any time $t > 0$.

Variance exhibits mean reversion at the rate ϱ, towards the level θ. The two scalar Wiener processes *are correlated* with correlation denoted $\breve{\rho}$. The leverage effect would be incorporated by means of a negative correlation. k is a constant "skewness" coefficient. The reason for which the parameter k represents skewness is that when the volatility is high, the effect that a large negative value of the shock dw_2^* produces on the volatility is high as well, while it cannot be high when volatility is low. So, it tends to make volatility movements asymmetric.

The *very special* case of this model in which the correlation of the Wiener processes is equal to 1 is consistent with a version of the special Cox-Ingersoll and Ross (1985a) general-equilibrium economic model. In that model, the premium on the risk of exogenous state variables is

$$\frac{1}{\frac{\partial J}{\partial W}} \mathrm{cov}\left[\frac{\partial J}{\partial W}, V\right],$$

where J is the indirect utility function of the representative investor, and W his wealth. With a log-utility specification, the market price of risk of the Cox-Ingersoll and Ross general-equilibrium is equal to

$$\kappa = \sqrt{V}.$$

The value for the risk premium is:

$$-\sqrt{V} \times k\sqrt{V} = -kV,$$

as supposed in SDE (7). So, Heston's model, unlike the Hull-and-White model does allow non zero risk pricing relative to the marginal utility of wealth.

But there is a cost: option pricing is more involved than in the Hull-White case, since the transition densities of the process S are not known. Because their characteristic functions can be obtained, the option price can be calculated by a Fourier inversion, as follows.[4]

The characteristic function of the distribution of the logarithm of $S(T)$ is defined as

$$f(x, y, t; \phi, T) \triangleq \mathbb{E}^*\left[(S(T))^{i\phi} \mid S(t) = x, V(t) = y\right], \tag{8}$$

[4]See also Chapter 13 in Dumas and Luciano (2017).

where i is the imaginary number and $\phi \in \mathbb{R}$. Note that the expected value is taken over the terminal stock price $S(T)$, under the risk-neutral joint probability of the underlying and its variance, started at x and y at time t.

If S and V follow an Heston model, using the fact that the characteristic function is a (local) martingale, we can write the PDE for it:

$$\begin{aligned} & rx\frac{\partial f}{\partial x}(x,y,t;\phi,T) + \frac{1}{2}yx^2\frac{\partial^2 f}{\partial x^2}(x,y,t;\phi,T) \\ & + [\varrho(\theta - y) - ky]\frac{\partial f}{\partial y}(x,y,t;\phi,T) + \frac{1}{2}k^2 y\frac{\partial^2 f}{\partial y^2}(x,y,t;\phi,T) \\ & + \frac{\partial^2 f}{\partial x \partial y}k\breve{\rho}yx + \frac{\partial f}{\partial t}(x,y,t;\phi,T) = 0, \end{aligned} \tag{9}$$

with the obvious boundary condition:

$$f(x,y,T;\phi,T) = x^{i\phi}. \tag{10}$$

One can guess a solution of the form:

$$f(x,y,t,\phi,T) = x^{i\phi}e^{C(T-t)+D(T-t)y}, \tag{11}$$

where the functions C and D depend on time to maturity, and the boundary condition is $C(0) = 0$, $D(0) = 0$. Let us verify that the guessed solution works. Substituting (11) into (9) gives

$$\begin{aligned} & ri\phi - \frac{1}{2}y\left(\phi^2 + i\phi\right) + [\varrho(\theta - y) - ky]D + \frac{1}{2}k^2 yD^2 + i\phi Dk\breve{\rho}y \\ & + \left[\frac{\partial C}{\partial t} + \frac{\partial D}{\partial t}y\right] = 0. \end{aligned}$$

If we collect the terms in y and distinguish them from the terms that do not depend on y, we get

$$\begin{cases} ri\phi + \varrho\theta D + \frac{\partial C}{\partial t}(T-t) = 0, \\ -\frac{1}{2}\left(\phi^2 + i\phi\right) - (\varrho + k)D + \frac{1}{2}k^2 D^2 + i\phi Dk\breve{\rho} + \frac{\partial D}{\partial t}(T-t) = 0. \end{cases} \tag{12}$$

These are two ODEs, for the functions C and D respectively, the first being linear, the second Riccati, which can be solved analytically.

The cumulative transition probability P^*, namely the cumulative probability function of $S(T)$ from 0 to $\mathcal{K}$, is then obtained by a kind of Fourier inversion:

$$P^*(x,y,t;\mathcal{K},T) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{\mathcal{K}^{-i\phi}f(x,y,t;\phi,T)}{i\phi}\right]d\phi, \tag{13}$$

where Re denotes the real part. From this cumulative transition probability, the option price is easily obtained, since the option payoff depends on the underlying and the price is an expectation over the cumulative probability function of the underlying.

By fitting as closely as possible the theoretical option prices to the observed ones, calibration of the model parameters, including the risk premium, can be obtained. This is the implied risk premium, corresponding to the specific risk-neutral measure dictated by market prices. Usually, the implied parameters are obtained from option quotes on a specific day (a single cross section of option prices). As a consequence, implied parameters tend to be somewhat unstable.

To sum up, the stochastic volatility models described so far present a number of complexities. First, they generate an incomplete market, with an infinity of possible risk premia, so that some arbitrary assumption must be made regarding the market price of volatility risk. We consider the assumption "arbitrary" because the market price of volatility risk is not observable and the one chosen may not be compatible with any economic-equilibrium specification. In theory, one could build an equilibrium model, but it would not be flexible enough to match a set of observed option prices. Second, the analysis in Dumas *et al.* (1998) pointed to the difficulty of eliciting a model for volatility, a quantity that is not directly observable. The difficulty was present in the case of DVF. It is even more present when volatility is assumed to follow a separate stochastic process. One is caught in a perilous loop: in order to figure out the right option pricing model, one must know how volatility behaves, and one is able to elicit volatility out of option prices only when one has calibrated the right volatility process.

The problem will be addressed by the more recent arbitrage-free formulations, which we develop in Section 2 below. Another, more pragmatic possibility is available in discrete time in the form of purely statistical models such as Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, which make it possible to circumvent the calibration difficulties at least. We do not study them here, but refer the reader to Dumas and Luciano (2017) for a comparison with the continuous-time models. Above all, one has to understand which modelling restrictions, if any, are needed in order to reap these calibration advantages.

2 Pricing a derivatives on the basis of the underlying and other derivatives

In the spirit of Black and Scholes, we can try to return to complete markets by pretending we know that the entire system of prices of options written on the *same underlying security* is driven by a K-dimensional Wiener process, so that we can utilize their initial values as an indication of their future behavior. In this Section, we are going to recall the basic call option price properties of Breeden and Litzenberger and the Dupire PDE, to define the so-called forward volatility of the underlying S, in terms of the call prices, and vice versa. We extend to stochastic volatility by showing that, using Itô's lemma, the relationship from the call prices to the volatility and its "inverse" have an analogue for their own volatilities.

2.1 The Breeden-and-Litzenberger hint

The whole theory of forward stochastic volatility relies on the link between call prices and the density of the underlying process at each point in time, or, if you wish, on the possibility to go from the derivative price to the distribution of the underlying at a specific point in time. We study the problem assuming that the underlying is a diffusion process. If one assumes, in particular, that the underlying is driven by a single Wiener process, there is a straightforward way to back out the density of the underlying price at T from the price of derivative securities. It is customary to refer to the probability density so obtained as being *implied* in the valuation of derivatives. We call it also the risk-neutral density, and denote it as p^*, its distribution being P^*.

Denote by $C(t, \mathcal{K}, T)$ the price at date t of a *European* call option with strike $\mathcal{K}$ and maturity date T on a given underlying security $S(t)$.[5] Without loss of generality, let us assume that the interest rate r is zero. From the fundamental theorem of asset pricing

$$C(t, \mathcal{K}, T) = \mathbb{E}^* \left\{ [S(T) - \mathcal{K}]^+ | S(t) = S \right\}.$$

Let us write the expectation which is the call price in terms of the distribution function of the underlying at time T, $P^*(x)$, conditional on the time-t price. For simplicity, set $t = 0$:

$$\begin{aligned} C(0, \mathcal{K}, T) &= \mathbb{E}^* \left\{ [S(T) - \mathcal{K}]^+ | S(0) = S_0 \right\} \\ &= \int_0^{+\infty} (x - \mathcal{K})^+ dP^*(x) \\ &= \int_{\mathcal{K}}^{+\infty} x dP^*(x) - \mathcal{K} \int_{\mathcal{K}}^{+\infty} dP^*(x) \\ &= \int_{\mathcal{K}}^{+\infty} x dP^*(x) - \mathcal{K} (1 - P^*(\mathcal{K})). \end{aligned} \tag{14}$$

If we now take the derivative twice with respect to $\mathcal{K}$, we get

$$\begin{aligned} \frac{\partial C}{\partial \mathcal{K}} &= -1 + P^*(\mathcal{K}), \\ \frac{\partial^2 C}{\partial \mathcal{K}^2} &= p^*(\mathcal{K}). \end{aligned}$$

By repeating the exercise when $\mathcal{K}$ runs over the support of S, we get the whole density of the underlying as

$$p^*(x) = \left. \frac{\partial^2 C}{\partial \mathcal{K}^2} \right|_{\mathcal{K} = x}$$

Breeden and Litzenberger (1978) first recognized this result, with reference to a diffusion with non-stochastic volatility. We are going to use a sophisticated version of it in the proof of Theorem 3.

[5] No similar theory exists for American-type options.

2.2 The Dupire forward PDE

The density p^* evolves according to the Fokker-Planck equation

$$\frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2 x^2 p^*\right) = \frac{\partial}{\partial T} p^*.$$

Note that, since $dP^* = p^* dx$, (14) implies

$$\frac{\partial C}{\partial T} = \int_{\mathcal{K}}^{+\infty} \frac{\partial p^*}{\partial T}\left(x - \mathcal{K}\right)^+ dx.$$

Using the Fokker-Planck equation, it becomes

$$\frac{\partial C}{\partial T} = \int_{\mathcal{K}}^{+\infty} \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2 x^2 p^*\right)\left(x - \mathcal{K}\right)^+ dx.$$

Integrating the right-hand side we get

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2 \mathcal{K}^2 \frac{\partial^2}{\partial \mathcal{K}^2} C, \tag{15}$$

which is a PDE for the unknown $C\left(t, \mathcal{K}, T\right)$ over $\mathcal{K} \geq 0$ and $T : t \leq T \leq \infty$ with *initial* condition $C\left(t, \mathcal{K}, t\right) = \left[S\left(t\right) - \mathcal{K}\right]^+$. The solution is a function $C\left(t, \mathcal{K}, T\right)$ of $\mathcal{K}$ and T. Equation (15) is called the "forward equation" of Dupire (1994).

2.3 Static arbitrage restrictions on call prices

As a function of strike and maturity date, the option prices $C\left(t, \mathcal{K}, T\right)$ form a *surface*. The underlying security is a particular element of that large vector or surface; its price is $S\left(t\right) \equiv C\left(t, 0, T\right)$, which does not depend on T. Let us specify directly the behavior of prices under some risk-neutral probability measure P^*. Under P^*, we know that the drift of the prices of all traded securities is the rate of interest r, which we still normalize to 0. The diffusion matrix of the options remains to be chosen.

At time t, for all $\mathcal{K}, T > t$, we can write:

$$\frac{dC\left(t, \mathcal{K}, T\right)}{C\left(t, \mathcal{K}, T\right)} = \sigma_C\left(t, \mathcal{K}, T\right) dw^*\left(t\right), \tag{16}$$

$$C\left(0, \mathcal{K}, T\right) = C_0,$$

$$C\left(t, \mathcal{K}, t\right) = \left[S\left(t\right) - \mathcal{K}\right]^+, \tag{17}$$

and, in particular, for $\mathcal{K} = 0$:

$$\frac{dS\left(t\right)}{S\left(t\right)} = \sigma_S\left(t\right) dw^*\left(t\right), \tag{18}$$

$$S\left(0\right) = S_0,$$

where $S_0, C_0 \in \mathbb{R}^+$, $w^*\left(t\right)$ is a K-dimensional Wiener process under the risk neutral measure and $\sigma_C\left(t, \mathcal{K}, T\right)$ contains K columns. The prices $C\left(t, \mathcal{K}, T\right)$

being prices of options written on the same underlying, they are restricted to satisfy some joint terminal and initial conditions and we need to impose some restrictions on $\sigma_C(t, \mathcal{K}, T)$. Otherwise, arbitrage opportunities would develop.

A number of static arbitrage restrictions on C were spelled out by Merton (1973). For instance, an option cannot be worth more than its underlying security. They all derive from the payoff conditions (17). It follows from them that:

- $[S(t) - \mathcal{K}]^+ \leq C(t, \mathcal{K}, T) \leq S(t)$
- $C(t, \mathcal{K}, T)$ is increasing in T, non-increasing and convex in $\mathcal{K}$
- $\lim_{\mathcal{K}\to\infty} C(t, \mathcal{K}, T) = 0$
- $\lim_{\mathcal{K}\to 0} C(t, \mathcal{K}, T) = S(t)$

Obviously, if we choose $\sigma_C(t, \mathcal{K}, T)$ in (16) arbitrarily, at some point one of these barriers or conditions will be breached. We need a way to specify the diffusion of C, or the underlying asset, in such a way that that does not happen. To do so, we define forward variance.

2.4 From the option surface to the forward variance and vice-versa

Given an option price function or surface $C(t, \mathcal{K}, T)$ at a given point in time t, for varying $\mathcal{K}, T > t$, on an underlying $S(t)$, the forward volatility is defined as:[6]

Definition 1 *The forward variance of $S(T)$ that obtains from the surface C is defined as*

$$v(t, \mathcal{K}, T) \triangleq \frac{2\frac{\partial}{\partial T} C(t, \mathcal{K}, T)}{\mathcal{K}^2 \frac{\partial^2}{\partial \mathcal{K}^2} C(t, \mathcal{K}, T)}. \tag{19}$$

In the limit $\mathcal{K} \to 0$,

$$v(t, 0, t) = \sigma_S(t).$$

This definition gives the volatility as a functional of call prices. If the static arbitrage conditions are satisfied, the result is positive:

$$v(t, \mathcal{K}, T) > 0 \text{ for } T \geq t, \mathcal{K} \geq 0. \tag{20}$$

In the reverse, if we give ourselves a forward-variance surface $v(t, \mathcal{K}, T)$, we can obtain the call prices solving the Dupire PDE:

$$\frac{1}{2} v(t, \mathcal{K}, T) \mathcal{K}^2 \frac{\partial^2}{\partial \mathcal{K}^2} C(t, \mathcal{K}, T) = \frac{\partial}{\partial T} C(t, \mathcal{K}, T), \tag{21}$$

[6] Evidently, since in practice options are traded only for a finite number of maturities and strikes, in any specific practical circumstance, the surface available is one for a discrete grid only. Interpolation methods can be used to fill in the gaps.

for the unknown $C(t, \mathcal{K}, T)$ over $\mathcal{K}$ and T, with *initial* condition (17). Let us give a name to that mapping:[7]

$$\text{Dupire: } v(t, \mathcal{K}, T) \rightarrow C(t, \mathcal{K}, T) \text{ for fixed } t.$$

The procedure is analogous to the recovery of bond prices from forward rates by the integral. It is less straightforward than an integral but it is equally well defined, as, under technical conditions, there is a unique solution to that Cauchy problem.

2.5 Implications for the variance of forward variance and options

Let us add one more layer to that, and think of variance as a diffusive process, with a (instantaneous) variance $\sigma_v(t, \mathcal{K}, T)$. We can define the diffusion coefficient of forward variance from the diffusion coefficient of option prices and vice-versa. Applying Itô's lemma to the definition (19), with K Wiener processes, the diffusion of v is

$$\sigma_v(t, \mathcal{K}, T) = 2\frac{\frac{\partial}{\partial T}\sigma_C(t, \mathcal{K}, T)}{\mathcal{K}^2 \frac{\partial^2}{\partial \mathcal{K}^2} C(t, \mathcal{K}, T)} - 2\frac{\frac{\partial}{\partial T} C(t, \mathcal{K}, T)}{\mathcal{K}^2 \left(\frac{\partial^2}{\partial \mathcal{K}^2} C(t, \mathcal{K}, T)\right)^2} \frac{\partial^2}{\partial \mathcal{K}^2}\sigma_C(t, \mathcal{K}, T). \tag{22}$$

In the reverse, suppose that we give ourselves the function $v(t, \mathcal{K}, T)$ and the diffusion $\sigma_v(t, \mathcal{K}, T)$ of v, can we get the diffusion of the options $\sigma_C(t, \mathcal{K}, T)$? The answer is positive, as we demonstrate now. Note that, given the function $v(t, \mathcal{K}, T)$, the Dupire mapping is well defined and can be performed *at every fixed point in time* t independently of the behavior of v over time. For that reason, we can apply Itô's lemma to the terms of Equation (21) to get

$$\begin{aligned} \frac{1}{2}\sigma_v(t, \mathcal{K}, T)\mathcal{K}^2 \frac{\partial^2}{\partial \mathcal{K}^2} C(t, \mathcal{K}, T) + \frac{1}{2} v(t, \mathcal{K}, T)\mathcal{K}^2 \frac{\partial^2}{\partial \mathcal{K}^2}\sigma_C(t, \mathcal{K}, T) \\ = \frac{\partial}{\partial T}\sigma_C(t, \mathcal{K}, T). \end{aligned} \tag{23}$$

For fixed t, the function $\sigma_v(t, \mathcal{K}, T)$ being given and the function $C(t, \mathcal{K}, T)$ being obtainable from the Dupire mapping, we can solve that PDE for the function $\sigma_C(t, \mathcal{K}, T)$ over $\mathcal{K} \geq 0$ and $T : t \leq T \leq \infty$ with *initial* condition

$$\sigma_C(t, \mathcal{K}, t) = \begin{cases} 0 & \text{if } \mathcal{K} > S, \\ \sigma_S(t) & \text{otherwise.} \end{cases}$$

Let us give a name to that mapping:

$$\text{DDupire: } \sigma_v(t, \mathcal{K}, T) \rightarrow \sigma_C(t, \mathcal{K}, T) \text{ for fixed } t.$$

Proposition 2 *The function $\sigma_C(t, \mathcal{K}, T)$ so defined satisfies the static arbitrage restrictions automatically.*

[7] The mapping is not defined for $\mathcal{K} = 0$. It needs to be extended by taking the limit $\mathcal{K} \rightarrow 0$.

3 Forward variance as a forecast of future spot variance

The forward variance has a very telling interpretation, which justifies its name:

Theorem 3 *The forward variance is equal to the expected value at time t under the risk neutral measure of the variance of the underlying at time T conditional on the underlying price being equal to $\mathcal{K}$ at that time:*

$$v(t,\mathcal{K},T) = \mathbb{E}_t^* \left[\sigma_S^2(T) \mid S(T) = \mathcal{K}\right]. \tag{24}$$

Proof. We give the proof only for the case in which the underlying price $S(t)$ and its scalar volatility $\sigma_S(t)$ follow a joint Markov diffusion process

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \sigma_S(S,t)\, dw(t), \\ d\sigma_S(S,t) &= \mu_S(t)dt + \gamma(t)dZ(t) \end{aligned}$$

where dw and dZ are correlated Wiener processes with correlation ρ and with transition density $p^*(S(t), \sigma_S(t); S(T), \sigma_S(T), T-t)$ from time t to time T under the risk-neutral measure.[8] Remind that the rate of interest is equal to 0. The price of the option is

$$C(S,\sigma_S,\mathcal{K},T-t) = \int_{\mathcal{K}}^{\infty} \int_0^{\infty} p^*(S,\sigma_S;x,y,T-t)(x-\mathcal{K})\, dydx. \tag{25}$$

It is obvious that one can generalize to stochastic volatility the equation of Breeden and Litzenberger. Indeed, taking a second derivative of both sides of (25) gives

$$\frac{\partial^2}{\partial \mathcal{K}^2} C(S,\sigma_S,\mathcal{K},T-t) = \int_0^{\infty} p^*(S,\sigma_S;\mathcal{K},y,T-t)\, dy,$$

which is the risk-neutral probability density that $\mathcal{K} \leq S(T) < \mathcal{K}+d\mathcal{K}$. Similarly,

$$\frac{\partial}{\partial T} C(S,\sigma_S,\mathcal{K},T-t) = \int_{\mathcal{K}}^{\infty} \int_0^{\infty} \frac{\partial}{\partial T} p^*(S,\sigma_S;x,y,T-t)(x-\mathcal{K})\, dydx.$$

Now, consider that p^* must satisfy the Fokker-Planck equation in two dimen-

[8] This heuristic proof was given in an unpublished note by Dumas (1995) and is a generalization of the appendix in Derman and Kani (1994). See also Carr and Madan (1998) and Britten-Jones and Neuberger (2000). A general and rigorous proof for semimartingales is to be found in Carmona and Nadtochiy (2009).

sions:

$$
\begin{aligned}
-\frac{\partial}{\partial y}\left[\mu_S \times p^*\left(S,\sigma_S;x,y,T-t\right)\right] \\
+\frac{\partial^2}{\partial x^2}\left[\frac{1}{2}y^2x^2p^*\left(S,\sigma_S;x,y,T-t\right)\right] \\
+\frac{\partial^2}{\partial y^2}\left[\frac{1}{2}\gamma^2p^*\left(S,\sigma_S;x,y,T-t\right)\right] \\
+\frac{\partial^2}{\partial x\partial y}\left[\rho\gamma x p^*\left(S,\sigma_S;x,y,T-t\right)\right] = \frac{\partial p^*}{\partial T}\left(S,\sigma_S;x,y,T-t\right),
\end{aligned}
\tag{26}
$$

subject to the boundary (or *initial*) condition:

$$
p^*\left(S,\sigma_S;x,y,0\right) = \delta_0\left(y-x\right). \tag{27}
$$

where $\delta_0\left(y-x\right)$ is the Dirac delta function at $x=y$. Multiply by $x-\mathcal{K}$ both sides of the Fokker-Planck equation and integrate with respect to x and y:

$$
\begin{aligned}
&-\int_0^\infty\int_{\mathcal{K}}^\infty \frac{\partial}{\partial y}\left[\mu_S \times p^*\left(S,\sigma_S;x,y,T-t\right)\right]\left(x-\mathcal{K}\right)dxdy \\
&+\int_0^\infty\int_{\mathcal{K}}^\infty \frac{\partial^2}{\partial x^2}\left[\frac{1}{2}y^2x^2p^*\left(S,\sigma_S;x,y,T-t\right)\right]\left(x-\mathcal{K}\right)dxdy \\
&+\int_0^\infty\int_{\mathcal{K}}^\infty \frac{\partial^2}{\partial y^2}\left[\frac{1}{2}\gamma^2p^*\left(S,\sigma_S;x,y,T-t\right)\right]\left(x-\mathcal{K}\right)dxdy \\
&+\int_0^\infty\int_{\mathcal{K}}^\infty \frac{\partial^2}{\partial x\partial y}\left[\rho xy\gamma p^*\left(S,\sigma_S;x,y,T-t\right)\right]\left(x-\mathcal{K}\right)dxdy \\
&=\int_0^\infty\int_{\mathcal{K}}^\infty \frac{\partial}{\partial T}p^*\left(S,\sigma_S;x,y,T-t\right)\left(x-\mathcal{K}\right)dxdy.
\end{aligned}
$$

Integrating by parts and/or integrating depending on the terms, and assuming that the transition density function p^* has the property that $p^*=0$ and $\partial p^*/\partial y = 0$ at $y=0$ and $y=\infty$, the first, third and fourth terms vanish and we are left with:

$$
\begin{aligned}
&\int_0^\infty \frac{1}{2}y^2\mathcal{K}^2p^*\left(S,\sigma_S;\mathcal{K},y,T-t\right)dy \\
=\;&\int_0^\infty\int_{\mathcal{K}}^\infty \frac{\partial}{\partial T}p^*\left(S,\sigma_S;x,y,T-t\right)\left(x-\mathcal{K}\right)dxdy.
\end{aligned}
$$

But the left-hand side is

$$
\frac{1}{2}\mathcal{K}^2\mathbb{E}_t^*\left[\sigma^2(T)\mid S(T)=\mathcal{K}\right]\times\frac{\partial^2}{\partial\mathcal{K}^2}C\left(S,\sigma_S,\mathcal{K},T-t\right),
$$

and the right-hand side is $\partial C/\partial T$ so that, comparing with Definition (19), we can identify the conditional expected value with the forward variance $v(t,\mathcal{K},T)$.
■

The name forward variances then was adopted because they are expectations of the future spot variance of the underlying, somewhat like forward interest rates are expectations of future instantaneous rates.

To conclude this Section, note that the forward variance is not the implied variance. To put it simply, while the implied variance is model-dependent, because it derives from Black-Scholes, the forward variance is "model free," as is the result of Breeden and Litzenberger. Also, note that we assumed that call prices for all strikes and maturities are quoted and observed. This makes the market *de facto* complete, and makes the measure P^* that is implicit in market prices, the unique equivalent measure of the market.

3.1 Summary of the option valuation procedure

Recall that the option price surface is driven by a Wiener process w of dimension K. To implement the procedure for option valuation, suppose that there exist K "reference" options all written on the same underlying security. This means that, under the effective measure, we observe the following processes:

$$\begin{aligned}\frac{dC\left(t, \mathcal{K}_i, T_i\right)}{C\left(t, \mathcal{K}_i, T_i\right)} &= \mu_C\left(t, \mathcal{K}_i, T_i\right) dt + \sigma_C\left(t, \mathcal{K}_i, T_i\right) dw\left(t\right), \\ i &= 1, ..., K,\end{aligned}$$

where the diffusion matrices of the K options are consistent with each other in the sense that

$$\sigma_v(t, \mathcal{K}, T) \quad \overset{\text{DDupire}}{\rightarrow} \quad \sigma_C\left(t, \mathcal{K}_i, T_i\right) \quad \text{for fixed } t \quad \forall i = 1, ..., K$$

for some common process $\sigma_v(t, \mathcal{K}, T)$.

Then we can define the market-price of risk vector (still assuming a zero rate of interest):

$$\kappa\left(t\right) \triangleq \sigma_K\left(t\right)^{-1} \mu_K\left(t\right),$$

where $\sigma_K\left(t\right) \triangleq \left\{\sigma_C\left(t, \mathcal{K}_i, T_i\right)\right\}_{i=1}^{K}$ and $\mu_K\left(t\right) \triangleq \left\{\mu_C\left(t, \mathcal{K}_i, T_i\right)\right\}_{i=1}^{K}$.

From κ we can define an exponential local martingale η as the real-valued process $\eta\left(t\right)$

$$\eta\left(t\right) \triangleq \exp\left\{-\int_0^t \kappa\left(s\right)^{\intercal} dw\left(s\right) - \frac{1}{2}\int_0^t \left\|\kappa\left(s\right)\right\|^2 \times ds\right\}. \tag{28}$$

The exponential martingale serves as a change of measure and allows us to price all options other than the K reference options as well as any derivative written on the same underlying, by computing the expectation of its discounted payoff. This means using the Fundamental theorem of asset pricing, or the so-called martingale pricing technique (see for instance Dumas and Luciano, 2017, Chapter 6).

Equivalently, we can define the process w^* that is a Wiener process under the risk neutral measure:

$$dw^*(t) = dw(t) + \kappa(t)\,dt.$$

change the process of the underlying using Girsanov's theorem and compute the value of any derivative starting from the changed process for the underlying.

Knowing κ, η and the risk neutral probability from the observation of the process of K reference options, one may, indeed, evaluate any derivative written on the same underlying security, including exotic options, simply by calculating a conditional expected value under the risk-neutral measure. For instance, the price of all options, beyond the K reference ones, follows.

Note also that the SDE on any option under the effective measure, and thereby their empirical expected rate of return, would be given by

$$\frac{dC(t,\mathcal{K},T)}{C(t,\mathcal{K},T)} = \sigma_C(t,\mathcal{K},T)\,\sigma_K(t)^{-1}\,\mu_K(t)\,dt + \sigma(t,\mathcal{K},T)\,dw(t)\,, \qquad \forall T$$

provided that

$$\sigma_v(t,\mathcal{K},T) \quad \overset{\text{DDupire}}{\rightarrow} \quad \sigma_C(t,\mathcal{K},T) \quad \forall \mathcal{K}, T$$

for a process $\sigma_v(t,\mathcal{K},T)$ that coincides at the points $\{\mathcal{K}_i, T_i\}_{i=1}^{K}$ with the one that was common to all the K reference options.

3.2 VIX

Using the conditional expectation of Theorem 3, one can obtain from current option prices an estimate of the market's *unconditional*, risk-neutral expectation of the future variance. One only has to condition down by multiplying the expression (19) by the conditional probability $\partial^2 C/\partial\mathcal{K}^2$ and integrating over $\mathcal{K}$:

$$\mathbb{E}_t^*\left[\sigma_S^2(T)\right] = 2\int_0^\infty \frac{1}{\mathcal{K}^2}\frac{\partial}{\partial T}C(t,\mathcal{K},T)\,d\mathcal{K}.$$

Suppose one wanted to obtain an average of this expected value over a future period of time $[t_1, t_2]$, with $t < t_1 < t_2$:

$$\int_{t_1}^{t_2} \mathbb{E}_t^*\left[\sigma_S^2(u)\right] du = 2\int_0^\infty \frac{1}{\mathcal{K}^2}\left[C(t,\mathcal{K},t_2) - C(t,\mathcal{K},t_1)\right] d\mathcal{K}. \qquad (29)$$

This says that the market price of a contract on future variance of the underlying can be obtained directly by integrating over their strike prices a continuum of differences between option prices of *identical strikes and different maturity dates*. This is the "price of a variance swap" as defined in Carr and Madan (1998) and Britten-Jones and Neuberger (2000).

The Chicago Board Options Exchange (CBOE) publishes daily an index called "the VIX", which has been much in the news during the 2008 crisis and in more recent years. It used to be calculated from the Black-Scholes formula as a mix of option prices, where the weight of each option depended on the volume traded. Since 2003, it is calculated according to formula (29). So, according to the left hand side of that expression, it is an average of expected spot volatilities. The right hand side says that it can be computed as an integral (a sum, when discretized) of the difference between the final and initial call prices, over all possible strikes, Under the new definition, VIX contracts could then be exactly replicated if a continuum - over strikes - of options were available. More on the replication is indeed in Carr and Madan (1998).

Buraschi and Jackwerth (2001) argue, using the VIX data, that volatility risk receives a non zero price (over and beyond consumption risk) in the market.

4 Conclusion

Lacking a rendition of the way information arrives to financial markets, it is not practical at this point to construct a structural model of volatility behavior. Information could be arriving smoothly and continuously or in big lumps at isolated times, with very different behavior for resulting volatility.

So, mostly as much as the behavior of future interest rates, at least in expectation, is captured by the forward rates implied in observed zero-coupon bond prices, future volatility is extracted from current prices. Using a premonitary hint of Breeden and Litzenberger, the prices to be used are not the ones of the underlying asset, but the ones of the derivatives. Of all the approaches reviewed in this chapter, the stochastic volatility implied in the current volatility surface is the most promising one. The forward variance is the expected value of the future one either under a forward measure or, equivalently, under the plain risk neutral measure but conditionally on the underlying price hitting a particular value at a particular time. This was the core of the current paper, together with its empirical consequences for option pricing and the VIX.

5 References

Breeden, D. T., and R. Litzenberger, 1978, Prices of State-Contingent Claims Implicit in Option Prices, *Journal of Business* 51, 4, 621-651

Britten-Jones, M. and A. Neuberger, 2000, Option Prices, Implied Price Processes, and Stochastic Volatility, *The Journal of Finance*, 55, 839-866.

Buraschi, A. and J. Jackwerth, 2001, The Price of a Smile: Hedging and Spanning in Option Markets, *Review of Financial Studies*, 14, 495-527.

Carmona, R. and S. Nadtochiy, 2009, Local Volatility Dynamic Models, *Finance and Stochastics*, 13, 1-48.

Carr, P., and D. Madan, 1998, Towards a Theory of Volatility Trading, in R. Jarrow, ed., *Volatility: New Estimation Techniques for Pricing Derivatives*, London: Risk Books, 417-427.

Cox, J.C., J. Ingersoll, and S. Ross, 1985a, An Intertemporal General Equilibrium Model of Asset Prices, *Econometrica* 53, 363-3854.

Cox, J.C., J. Ingersoll, and S. Ross, 1985b, A Theory of the Term Structure of Interest Rates, *Econometrica* 55, 385-407.

Derman, E. and I. Kani, 1994, Riding on a smile, *Risk*, 7, 32-39

Derman, E. and I. Kani, 1997, Stochastic implied trees: Arbitrage pricing with stochastic term and strike structure of volatility, *Quantitative Strategies Technical Notes*, Goldman Sachs, April.

Dumas, B., 1995, The Meaning of the Implicit Volatility Function in Case of Stochastic Volatility, unpublished note (meant to be an appendix to Dumas *et al.* (1998)) available at: http://faculty.insead.edu/bernard-dumas/research.

Dumas, B., J. Fleming. and R.E. Whaley, 1998, Implied Volatility Functions: Empirical Tests, *Journal of Finance*, 53, 2059-2106.

Dumas, B. and E. Luciano, 2017, *The Economics of Continuous-time Finance*, MIT Press.

Dupire, B., 1994, Pricing with a Smile, *Risk*, 7, 32-39.

Heath, D., R. Jarrow, and A. Morton, 1992, Bond Pricing and the Term Structure of Interest Rates: a New Methodlogy for Contingent Claims Valuation, *Econometrica* 60, 77-105.

Heston, S., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 327-343.

Hobson, D. G., and L. C. G. Rogers, 1998, Complete Models with Stochastic Volatility, *Mathematical Financ*e, 8, 27-48.

Hull, J., and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatilities, *The Journal of Finance* 42, 281-300.

Merton, R.C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141-183.

Platania, A. and L. C. G. Rogers, 2005, Putting the Hobson-Rogers Model to the Test, working paper, University of Padova.

Rubinstein, M., 1994, Implied binomial trees, *The Journal of Finance*, 49, 771-818.